IN THE UNITED STATES DISTRICT COURT
FOR THE EASTERN DISTRICT OF PENNSYLVANIA

- - -

IN RE:	:	2:12-md-02342-CMR
	:	PHILADELPHIA, PA
ZOLOFT	:	
(SERTRALINE HYDROCHLORIDE)	:	September 1, 2015
PRODUCTS LIABILITY LITIGATION	:	3:27PM

- - -

TRANSCRIPT OF DAUBERT HEARING DAY 2
BEFORE THE HONORABLE CYNTHIA M. RUFE
UNITED STATES DISTRICT JUDGE

APPEARANCES:

For the Plaintiff: JOSEPH ZONIES, ESQ.
REILLY POZNER LLP
1900 SIXTEENTH STREET, SUITE 170
DENVER, CO 80202
303-893-6100
jzonies@rplaw.com

For the Defendant: MARK S. CHEFFO, ESQ.
QUINN EMANUEL URQUHART & SULLIVAN
51 MADISON AVENUE, 22ND FLOOR
NEW YORK, NEW YORK 10010
212-849-7000
markcheffo@quinnemanuel.com

AUDIO OPERATOR: CSR OPERATOR
TRANSCRIBER: JANINE THOMAS
NOTARY PUBLIC

(Proceedings recorded by electronic sound recording, transcript produced by transcription service.)

VERITEXT NATIONAL COURT REPORTING COMPANY
MID-ATLANTIC REGION
1801 Market Street - Suite 1800
Philadelphia, Pennsylvania 19103
(888) 777-6690

I N D E X

Q. That new data while it was reported as a 1.0 when it includes the Alwan data that necessarily the new data that was not included in Alwan would've had a different effect?

A. Well, here it was a little different from Jimenez-Solem because they did provide in their supplement the ability to break out the data see the new data and directly compare it. And the new data showed a stronger, not very strong, but a stronger association than originally in Alwan. This is just I think for -- this is not for all cardiac this is just septal defects. And refuse was the only paper that I've seen that didn't actually talk about all cardiac and that's because Alwan didn't find it interesting and the original and they only focused on results from the earlier paper in this case.

Q. So, I put up there 1.2, 1.3, did you actually do that calculation at some point?

A. I did.

Q. And do you recall what it was?

A. My memory is it's 1.22, but I hate to go by memory.

Q. Dr. Jewell that -- when you're examining the associations your primary focus as you said was on Zoloft all cardiac outcomes and this is a [indiscernible] that captures those outcomes; is that right?

A. That is correct.

Q. And it has Cornum, Wemacher [ph] and it says next To Wemacher severe CHD. Did Wemacher also have a CHD finding

over a cardiac heart defect finding and a severe cardiac heart defect finding?

A. Yes. Wemacher's I think the only paper, there are other papers who talked about doing this which has said let's not look at all cardiac heart defects, because there's a concern about some of those are reported differentially if you're exposed or not, because of a variety of reasons we could discuss.

So, one way to get around detection bias is to focus on severe cases which are likely to be detected under all circumstances and.

THE COURT: I'm sorry.

THE WITNESS: That's okay. That's what Wemacher did. They did report had results for all cardiac defects and then they looked at the results for the severe. So, this slight that's just popped back again has the one with the severe CHD in it.

BY MR. ZONIES:

Q. And then you've also -- and I want to be clear Dr. Jewell, you reported all of the outcomes for all of the studies; right? You're not trying to duplicate and make it look like you're trying to slide one by on the judge; are you?

A. No. Again, it's important to pointer out that Cornum, Jimenez-Solem and Petersen are all there near the top. They're all Danish and there is some overlap. I don't want to

go over the territory again. So they performed less replication because you have to do the kind of discussion I just gave. You have to think about it a little bit more carefully.

And Reese Collen [ph] and Collen are in there at the bottom and they're, you know, Collen is really superseded by the Reese Collen they don't differ that much, but once you take that into account, this is all of the reports and all the literature that is important.

Q. And is there a reason you wouldn't want to look at all it and try to understand it?

A. No. I always want to look at it all.

Q. You also, Dr. Jewell report separately for septal [indiscernible]; is that right?

A. Yes, and here you can see the confidence in those getting whiter because, of course less, septal is a subcategory of all cardiac so you'd expect less precision, so that's what you see.

Q. Did you decide to create these all septal categories or this is what's provided in the scientific literature?

A. No. These are all provided in the original papers noted there.

Q. Dr. Jewell, when you're looking at the outcome of all cardiac and this is what you see, why do you both to look down at septal defects?

A. Well, I discussed this in my report and it's a conundrum that faces statisticians from an epidemiology from day one. You want to look at the association between an exposure and an outcome and if the outcome's pretty bad and some of these outcomes are bad, pretty bad or if the outcome is cancer that's pretty bad. It's fortunately often quite rare and these birth outcomes do not occur fortunately one out of every two births, otherwise, we'd have lots of them and we could get very, quite narrowly into specific subcategories, because there'd be a lot of them available for study that already all cardiac birth defects are not that common so you have to get a lot of woman and we saw numbers this morning, a million woman to get data. You have to look at a lot of pregnancies to even get close to some information that's reliable on a big category like all cardiac outcomes. That's why statisticians and epidemiologists have always done that of grouping outcomes that they think are reasonable, I wouldn't do it, I'm a statistician, but you get the expert when the first studies at Twenty Hills, first studies of lung cancer came out, we know there are different kinds of lung cancer, asbestos and smoking actually tend to raise the risk of slightly different kinds of lung cancer.

To you and I it's just lung cancer, but to the people who understand pathology of the lung there are different kinds of lung cancers, but when they did those studies, early

studies of smoking, because it was hard, an observational studies to get definitive information that would change policy they lumped them all together, because they needed sufficient data. And that choice is almost always there in every epidemiologic study, what's the outcome? How do you group them? Where -- if you go too big you have a chance of putting in [indiscernible] outcomes which have nothing to do with exposure and that will weaken your ability to see association.

On the other hand, so you get bias, basically and dilution. On the other hand you get very precise and say, I really think it's this specific heart defect where the problem is and you only get that specific one occurring one in 10,000 pregnancies you're going to end up with no data. So, how do the epidemiologist resolve that problem?

They get the experts together, they decide these bigger subcategories and they essentially draw that line of where they think it makes sense. Some of the authors in here have sufficiently small studies that they didn't think it made sense to go -- even to report subcategories and they had rules about that, but all of them had sufficient cases to report about all cardiac, as we've already discussed. That's the only reason I use that essentially as a category that made sense.

Now, do you want to look at the subcategories? Sure, suppose I'd seen a large ratio for all cardiac of two and then

you come along and say, I have a client and it's a cardiac defect, but it's this very specific one, a TGA, let's say, transposition of the great arteries. I say, oh, well, I'll go look at the data. There's hardly any data reported. It occurs very rarely. I think only one paper reported separately on TGAs, there's no information. Are we lost? No.

Suppose there had been data on TGA as a subcategory, remember, I said all cardiac idea seen in a [indiscernible] of two, but in TGA I saw consistently odds ratio .9. Well that would tell the statistician even though in the big category it looks like something is going on, it doesn't seems to be mirrored in the subcategory. The subcategory seems to be something that's in there, but there doesn't seem to be anything going on.

So a statistician will want to look at the big category to get precision and then look at the pattern of the relationship in subcategories to see if there's any, that seem to be pointing in a quite different direction and that's why I looked at the subcategories to the extent they were reported.

Q. So, Dr. Jewell, let's take a look at that. This is your cardiac outcomes with outcomes of 3, 2.8, 2.7, 2.36, 1.74, 1.5 and as you start to go into the subcategories you still see in septal consistently increased risk; right?

A. Yes.

Q. And if you look at reported differently here, above the

line would be increased risk, below the line is decreased risk for ASD as you get more refined atrial septal defects, ventricular septal defects, ABSDs, are you seeing any sort of a consistent patter even if you go to the other cardiac de fact which captures hypoplastic left heart, contrology of [indiscernible] RVOT, LVOT, all above this line positive associations. What does that tell you about the heterogeneity or not of these various cardiac outcomes?

A. Well, to do it justice, of course you have to spend some of the hours spent looking at it that in brief if you think of those flores plot [ph] flashing before your eyes all cardiac then the subcategory of septal then, a subcategory of let's say atrial septal defects or ventricular septal defects, the pattern of the flores plot looks the same as you go down. In other words what I'm looking for is there's a subcategory here which doesn't show any increase in risk systematically. It looks different from the big flores plot and that isn't apparent in this case, but that's why we look at it.

You have to recognize when you go down to the subcategories everything is more variable because there's less information, but even with that where you'd expect the results to bounce around a bit, I think you see enough here to see that there isn't a subcategory here that's demonstrably showing a different pattern of association than what you saw in the catchall all cardiac subgroup.

Q. Dr. Jewell, when you, after you examined the associations for pulling them out and looking for these generalizable thoughts about the subcategories did you then assess the -- in your methodology assess these associations for whether or not they were simply the product of chance ren variation?

A. Yes, there's been many times in my life where that's, when the story is stopped, because the data just doesn't look like it's sufficiently strong to be anything other than chance and that happens.

Q. When we're looking at these associations are a function of for example ren [ph] variation or systematic variation can you help us understand what that means, in other words, this is a 3.6 outcome here, so would bias and confounding if it were in play to make that disappear how big of an effect would you be looking for?

A. Well, we can quantify that, you would have to have a very, another risk factor confounding factor for example that would be as strongly related to the outcome as the 3s mentioned there to completely wipe that out.

Q. So, something going on in confounding or bias would have to be three times more prevalent if that's the rights to the -- I knew you'd say that. So, it would have to occur, have a three times impact on the odds ratio.

A. Do you want to restate that question before I give you

a grade?

Q. Yes. Could you restate the --

A. Yes. For confounding, we know that -- let's just take an odds ratio of three, if that is explainable by a confounding, so the true odds ratio is one, once you adjust for this hypothetical confounding factor. That hypothetical factor in itself has to have an odds ratio of at least 3. And it also has to have a relationship to, in this case Zoloft that's the exposure you're interested in. That's also a strong as three, so there isn't just, you can't just sprinkle a little confounding and wipe out the three, that's was the argument behind Bradford Hills first criteria, the strength of association.

It's -- when we saw an odds ratio of 10 for lung cancer and smoking the arguments of confounding went away and there were very, very strong arguments of confounding presented not only by industry at that point, but also by preeminent statisticians that thought smoking and lung cancer was just confounded, there was nothing to do with smoking, but when the size of it became apparent by Hill's work and by Dull's [ph] work then it became impossible to explain by some unknown confounder that we occasionally see that language, Mr. Cheffo put it out, but it's an unknown confounder.

Well, that's not, it's not so unknown it's got to be, it's a pretty strange confounder if it's that big and it's

still hiding from us.
Q. So, Dr. Jewell when assessing chance it is and I think you've said important, one of the factors that's important is statistically significant results.
A. Well, if there's nothing that's significant at all anywhere and all these results are just -- if that's the plot of all the associations they've all been sort of some to the right of one, some to the left of one, I would have just told you my report and you probably wouldn't have hired me, but that would've been the truth and that's where it would've gone.

So, if there was nothing going on, there's nothing going on. And that just happened to me probably more times than I found something going on, it's just part of life. But -- so that's the first sort of sign of [indiscernible] it seems to me. Let's not waste your time and my time if it's apparent to me that there's really no evidence one way or the other.
Q. And here did you find that there was evidence?
A. Yes.
Q. And one of the things that I discussed in my opening was Bonferroni context, is this something that you actually support doing to test this question?
A. I do support it in the context of the individual studies. If you take an individual study and you go fishing

for associations or what's the word that was used a little bit -- torturing, if you torture the data by looking at every possible association, gee, if you have red flowers in your front yard, you tend to have a heart attack and you find that association. If you look for enough you'll find them, that's what multiple comparisons are about, because it's like double jeopardy, it's triple, you're just trying -- you're just random -- you're sort of picking at straws, clutching as straws to find an association in a single study. That is a fallacy and we recognize it, we're particularly sensitive of it now with genetic studies in a single population, because there's 30,000 genes to look at to see if they're associated one by one with an outcome.

So, you have to have very string stringent controls then although, it's like kind of what I would call fishing for associations there. And Bonferroni was a way to adjust the P value to make those stringent criteria within a single study. And you can do it here Jimenez-Solem is the one major study which survives that kind of criterion. The others don't, big surprise, because we've already said time and time again these studies individually find it very hard to be definitive because there isn't a huge amount of data, so you wouldn't expect single studies really to show significance necessarily always which we've seen they don't, let alone surviving even stronger criteria, it doesn't make sense, and in fact, when

you go to looking at multiple studies the issue of multiple comparisons essentially dissipates for reasons I outlined in the report.

Q. And you've reviewed some of the e-mails I'm not going to go through these in detail, between other epidemiologists who were involved in these types of studies including Dr. Mitchell, Dr. Hernandez-Diaz, Dr. Greenland discussing the importance of significance or not related to a P value. Do you generally agree with these statements about slavish worship to that dichotomy?

A. Yes, I think in fact all the e-mails that I saw about this issue we all agreed accept for the editor at the New England Journal who agreed but said the rules are I've got to do something different. And then there's little e-mails about how could this editor believe this is crazy, but still make us do it, but, yes, but avenue statistician and epidemiologist in those e-mails agreed entirely that the whole thing was not relevant really.

Q. If you limited your inquiry in this case, Dr. Jewell, solely to the statistically significant results, would there be any results to the left of one?

A. No. As I said in my report there are some studies that show a negative association where the odds ratio comes in less than one, so the risk seems a little bit lower in woman exposed to Zoloft, but none of those achieved statistical

significance. And in some sense the slight here points out why one shouldn't limit to significance and if I had, how I would've been excoriated by everyone in sight, because this is all I would've talked about were these studies and everyone would've been unhappy. You would've been unhappy for different reasons than Mr. Cheffo, but I think both of you would've been unhappy with me, I wouldn't have done my job because you shouldn't and that just tells me that you know already you've got to look at all the evidence, because in the law you always look at all the evidence you don't just look at the significant clues, you look at all the evidence. And some of it isn't as important as others and there's no question that's true some of these studies are not as important as others, but you got to look at it all.

Q. Now, there's a continuing, it's here in the [indiscernible] by Pfizer for the past years, Dr. Jewell that apparently you believe an outcome like this one in -- Alwan, that you believe an outcome like this in Alwan the bottom last one that is to the left of one and crosses one on the upper bound somehow shows an increased risk; do you believe that?

A. No. I was quite explicit in my report that that study shows a decreased risk, it's inconsistent with my opinions. There's no question about that.

Q. That study result right there is totally inconsistent with what you ultimately proved.

A. It's below one, that's correct. Though you can see what I do point out in my work how precise is that known, because that's going to be important. In that one you can see specifically is very imprecise. The confidence interval goes from .1 to I can't really read it because the scale is never two or something, so it's very imprecise, but it does not support causation in this case, there's no question about that.

Q. Now, is there something about this upper bound or some portion of this upper bound that you do at least talk about being not incompatible?

A. Well, if you take the one, let's go to the first one that sort of drops below one if you can point to that, that one right there. Yes, so there's one that's .9 something, is an odds ratio and the data says there's less risk in the Zoloft exposed women in that study. So, I can't use that to say Zoloft causes a birth defect, that would be crazy; right?

Q. I think so?

A. That would be crazy, somebody said nuts. I'd be nuts. But look at the length of that confidence interval. Now, I know it goes way down to the left, that's no one in this courtroom believes that Zoloft has an odds ratio of .2 or .3, but you're 5 times less likely to have a cardiac. In fact Pfizer doesn't believe that, they'd be selling the drug to protect you from a heart defect anymore than people believe

the very right end of the confidence interval reaches five or six.

So, no one believes those end points, I understand that, but if you look at the imprecision in that study while it does not support my conclusions it doesn't really contradict them in and of itself because it's quite compatible with an odds ratio of and there wasn't enough data in that study to precisely nail down that two was ridiculous. So that's all I'm saying. I'm not saying supportive, I'm just saying it doesn't flatly contradict that and that's true of many of those that are down there as you can see half of them down there below one the confidence band on the right goes quite far up into the range of increased risk.

I mean if you would just go back to that slide for a second, I don't want to prolong things, but if you said to me, can you write a report that says the odds ratio for Zoloft and cardiac defects is .3, do we have evidence that there's a three-fold decrease in risk?

So, here's the data .3 if you can run your pointer up there, cuts may be 2 at most, maybe it gets that third one. There's no one in this room would believe that. Even though there's no -- a few of them are significantly different from .3, some of them aren't as we just said, but the preponderance of the evidence makes it clear to everyone.

Now, we're not in the quite as crystal clear situation

when we're talking about one here, but that's the way you've got to do it and in Hill's language, you have to say what is the reason for this pattern? It's clearly not symmetric. What is the reason here and is there -- here's his language, is there any other way of explaining the asymmetry of that plot? Or any other answer equally or more likely than cause or effect and that's what the rest of the report goes through. Is there another reason why we're seeing these studies slanted to a positive association.

Q. And that's what you get into next, right, is systematic variation --

A. Right.

Q. Bias or confounding? In your own words you don't stop there and say, yes, well, jeez it was all -- looks like there was more on a risk, but you actually go in and then say is something else causing that?

A. Yes, and many of the authors have raised in their limitations of their work, some of the biases that can -- may have distorted what we're see, and so we've got to pay attention to that.

Q. And one of the main ones -- you looked at many in your report; is that right?

A. I did. I did and confounding in general do women Zoloft smoke more or are they more obese, these other risk factors for cardiac of birth defects, I want to point out,

just before that was very important to look at that all the authors tried their best, some more successful than others. The flores plot we just had all of the results adjusted for all confounding that those authors did.

These are not the crude results. These are already de-biased in some sense by getting rid of all the confounding that the original authors were able to do.

Q. Well, and to put a fine point on that Dr. Jewell, the Huybrechts' number for all cardiac is I think 1.09 here.

A. Yes.

Q. And not significant. If you had chosen to try to get one by on everybody here you actually would've put in the Huybrechts' crude which would be 1.27 and it would actually be up here as statistically significant; right?

A. That's correct. So it moved on. And that was adjusting for confounding not only by a lot of these factors like smoking it was also adjusting for this confounding by indication by adjusting for whether the women had a diagnosis of depression or not.

So, I just want to be clear that flores plot has taken out as much as I can of those biases to the extent they were reported by the authors.

Q. So, the Huybrechts for example would've gone from 1.27 statistically significant in its crude down to its 1.16 when they do the depression restricted, but you even reported the

propensities for 1.09.
A. Yes and that was just adjusting for other factors that the women -- other, many, many other factors that the women had to try and remove their impact on the question.
Q. So, had we chosen to put the crudes in here for example, you could imagine this whole thing would actually --
A. It would've shifted to the right. Not always, but by in large.
Q. Now, confounded by indication I'm not going to go through each of these Dr. Jewell, because again, these are in your report that there were no studies demonstrating that depression in and of itself causes cardiac malformations.
And then the other various techniques that the authors have used to determine whether or not it's depression including looking at other antidepressants, this is all in your report; correct?
A. It's all in my report, yes. I looked at that careful.
Q. But, you -- and Jimenez-Solem, this is also in your report attempted to do this through this pause group analysis; is that right?
A. Correct.
Q. And you've mentioned that the Knutsen [ph] article, I believe in your report which discusses why that pause group may actually have this increased risk. Could you talk about your thoughts on the Jimenez-Solem pause group, first of all

about the fact that it's all SSRIs and then what else you think of that analysis?

A. I actually don't believe I did comment on Knutsen, Yen about the paper at the time I wrote the report, but I wrote the same argument that Knutsen actually wrote about later and that was that the definition of being paused in Jimenez-Solem was that you had to be taking an SSRI up to a certain time period several months before conception, at which point you did -- had no more prescriptions, so you paused.

And then you had to pick it up again, after a certain -- within a time period after the pregnancy had ended. There was a time interval at the end. And I said in my original report that gave me some concern, because of the fact that that might introduce a bias that women who suffered any form of birth defect or any complication in their pregnancy might be more likely to go back on antidepressants and that that would bias the results, of course, of Jimenez-Solem and that's the same reason that Knutsen independently wrote about that I didn't know when I wrote the report.

Q. And how would that bias the results? Why would that make it seem like there's more risk in that pause group?

A. Because it means that you're more likely to get into the pause group if you've had a heart -- not a heart -- any defect. So, we'll raise, artificially raise the rates of defects in the group. And they, Jimenez-Solem points out and

it was pointed out this morning by Mr. Cheffo that this pause group had an increased risk. And that's one possible explanation, not necessarily of the entirety of the increased risk, but why it may be overstating the risk of a true pause group, because you've got this bias and the way you've defined the pause group.

Q. And the fact that it's looking and only reporting on all SSRIs does that give you pause?

A. Yes, because at the beginning of my report and what we've discussed earlier I was at pains to try and use Zoloft specific information unfortunately Jimenez-Solem pause group is a group of women who took any SSRIs, not women who took Zoloft and then paused. Now, there had to be some in there who were Zoloft, it would be rather fewer, because we do know, from Jimenez-Solem the fraction of the women exposed to SSRIs before pregnancy who were Zoloft specific, but they don't report in the pause group how many were Paxil, pause Paxil users, pause Paxil users or pause Zoloft users.

So I couldn't compare the pause Zoloft users to the non-pause Zoloft users. I did have the data and Jimenez-Solem talks about the data for Zoloft-exposed women who continued to you it, but when you compare it back to a pause group it's not Zoloft paused only, it's a mixture of other women.

Q. And is that in an analysis that given the data you had you thought would be of interest?

A. Well, the whole point of Jimenez-Solem was to try and pull out this pause group and see if they were at an increased risk and we've talked about why they might be artifactually because of this definition of you had to go back on the SSRI and Zoloft after the pregnancy. Nevertheless this was at the core of Jimenez-Solem's argument about confounding by indication.

The argument I think I'm not -- I'm using their words as best I can. The argument is that if the pause group is still at increased risk compared to unmedicated depressed or unmedicated women, depressed or not it can't be the drug, because they're not taking it, they paused. That was the whole beauty of the Jimenez-Solem attack on the question. It was creative, they had pharmacy records, they had the ability to get the timing from pharmacy records. So, that was the whole point.

Unfortunately in my view they did that, they looked at the pause and they said they're at increased risk, but they didn't complete the loop which is yes, but are they at increased risk compared to the ones who didn't stop?

If you really had an ideal experiment, I forget the nice word Mr. Cheffo used exquisite experiment, you would take a group of women, you want it, the trouble plaguing this field is who do you compare Zoloft users, pregnant Zoloft users to? And initially they would compare them to any women having a

pregnancy and then people said, oh, wait a minute the women who suffer from depression aren't really comparable. You're comparing an apple and orange so, then, you know, then papers like Huybrechts, they try to restrict that we're going to look at depressed women some who take Zoloft and some that don't. But still a little bit -- Zoloft users are not necessarily the same Paxil users or they're not the same as somebody using another antidepressant or someone who's depressed that isn't taking a drug.

So, the goal of an epidemiologist is to try and get how do you pick those controls? The comparison group? What you would ideally like is a group of Zoloft users to a pregnancy that's your exposed group and then have another group of women who are identical in all regards who didn't take Zoloft, everything else would be the same.

The only way we know how to achieve that methodologically is to randomize. So, you will take Zoloft users and you'd randomly pick half of them when they conceived and said you can't take Zoloft anymore.

If we had that experiment that would be exquisite, because there could be no difference between the groups except that they didn't get -- one half of them didn't get exposed. We can't do that and people have judged that's not ethical for obvious reasons. The closest thing in all of these studies to getting to that exquisite experiment that all statistician and

epidemiologists want to do is this paused analysis. And you want to therefore say, okay, these women were identical going into conception, they were all taking Zoloft. Some of them stopped taking it and now we have this ability to compare those who continued and those that don't and see what the results. And Jimenez-Solem didn't actually do that, they compared both groups to the women not taking any drug, but when you do the direct comparison you get the 60 percent increase in risk which can't be purely then, so the confounding by indication hasn't completely eradicated this fact that women who stayed on it suffered for more cardiac defects than women who didn't. And here's the data. There's that 1.6, 1.59 to be precise.

Q. Your Honor, I'm about to go into a new area.

THE COURT: Nice area to take a break. Let's take ten minutes.

(Whereupon there was a recess in the proceeding from 4:01 p.m. to 4:17 p.m.)

DEPUTY CLERK: All rise. Court is now in session The Honorable Cynthia Rufe now presiding.

THE COURT: Everyone please be seated.

MR. ZONIES: Your Honor, I'm not sure what you want to do in timing. I think I probably have in truth 40, 50 or so to go. Dr. Jewell is here with us tomorrow as well, so.

THE COURT: Certainly, but if you can complete

your direct it would be best. We think we should keep going as long as we can all refrain from dropping our water bottles on the floor. Noise, that was me. All right. Thank you.

MR. ZONIES: We'll look to complete his direct tonight, Your Honor, today, whatever.

THE COURT: I do need though if you're going to keep using the -- we can't find this exhibit in the exhibits and it's very important to your case. We just need a copy that needs to be enlarged. Let's just say it, we need bigger letters because it's very hard to follow on any of our screens. I'm sorry that it just can't be enlarged enough to see it all at once; all right.

Mr. ZONIES: Absolutely.

THE COURT: So maybe overnight somebody can collie and enlarge it for us.

Mr. ZONIES: We can do that.

THE COURT: Thank you. All right. Please proceed.

BY Mr. ZONIES:

Q. And Dr. Jewell, I just want to while we're talking about this flores plot --

THE COURT: That one.

THE WITNESS: That's the one that's hard to see.

BY MR. ZONIES:

Q. What it -- in your methodology Dr. Jewell, where does

this fit in?
A. Well, this is really in my report a visual aid just to be able to try and put on one page, as difficult as it is to read all of the results. This is a few more than in my report, because of the new studies. So this is just a visual aid to be able to refer to it to see all in one page the whole, all the results whether in this case it's for all of them, but then the other ones for subcategories. It's just helpful as a visual device, I'm not trying to infer from this plot something quantitative. I've only referred to it this afternoon in the asymmetry of the plot which I think is apparent to everyone in the room, but other than that I'm not trying to infer anything directly from it and I've talked about why doing a matter analysis of all of those is fraught with danger, very appealing, but fraught with danger and we can talk more about that and where one might go given the fact that these studies are actually not all on equal footing, and two, show such heterogeneity.
Q. So, if for example Dr. Jewell is this something where you plotted these results out the first week you started working on this and you looked at this and tried to decide something today or is this something that you do to help guys like me when you're presenting your opinions?
A. Actually I produced this or a plot like this at the very end of writing the report. I didn't actually have it, I

had enough little notebooks and things laying around then at the end when I got rid of all those I are thought this -- I need something to be able to refer to and I put this in at the end of the report to sort of summarize where I thought it might be useful.

Q. We were talking about confounding by indication and the exquisiteness of Jimenez-Solem's pause group and some of the shortcomings of that and now idea like to talk, Dr. Jewell, about the Huybrechts analysis and in particular a more exquisite look at this data; okay?

A. Correct. Okay.

Q. There's a slide on the screen now that is the Huybrechts 2014 study entitled Antidepressant Use in Pregnancy and the Risk of Cardiac Defects. Is this a study that you've spent a great deal of time reviewing?

A. Yes.

Q. How did and as it shows on the screen, how did Huybrechts in this 2014 study define whether or not women were considered to be exposed to an antidepressant and in particular Zoloft during the critical period of early pregnancy?

A. Yes, so, the Huybrechts team had Medicaid data on pregnancies and then they had to merge that data with data on pharmacy records associated with Medicaid, so they had -- this was not -- the data set they had to construct basically by

merging data sets, once they had done that to the extent, that was done well and effectively, they then had to define what they meant by exposure, because people go on and off prescriptions all the time and so the piece that's highlighted in there with regard SSRIs they looked at the most recent, if you will, the most recent prescription of an SSRI prior to conception and they could see the number of pills that would be in the prescription and count the days and then say, if the prescription if taken as prescribed would take the woman into taking those pills during postconception count that woman as exposed.

Q. And so they looked at supply of in this case Zoloft over a period of weeks before conception and tried to determine if that supply would allow a woman to have the pills in a medicine cabinet essentially during the critical period of pregnancy?

A. Right. They did not know if the woman actually took the pills pregnancy for any person for obvious reasons, that's not recorded in the database and that's of course a weakness with pharmacy records because we know that people don't always take the medication, I bet you've not always taken your medication when prescribed. So --

Q. I don't need medication.

A. It's a problem, but that's the best her team could do. And it's somewhat similar to Jimenez-Solem had similar

pharmacy records.
Q. So, Huybrechts initial analysis that we've discussed compared, I'm showing you a slide Dr. Jewell of Huybrechts analysis with an odds ratio of 1.27 that is significant at the P value .05. Can you describe what this slide shows?
A. Yes, once she's classified, once her team has classified the women and the pregnancies into exposed that's the red here and unexposed and in this case removing women who were exposed to other antidepressant, so they have no antidepressant exposure at this point during their pregnancy. Then she compared the rate of cardiac defects and other defects in the red group versus the blue group, not surprisingly the blue group is huge, it swamps the red group, but nevertheless there's a lot of pregnant women exposed to Zoloft in their database 14,040 and the rates are compared and they're 27 percent higher in the red group than the blue group.
Q. And is that a true crude odds ratio is that what that would be?
A. That's a crude odds ratio. Just look at the rate of cardiac defects that are also now in this merged data set.
Q. And then the Huybrechts authors did a what's called on this next slide, Huybrechts analysis depression restricted comparison. Can you describe for us please what they did for depression restricted and/or and what the results were?

A. So to get around the idea that it's depression itself that may be contributing to a rate of defects in the exposed women and not the control in a previously light where include women who had no diagnosis of depression, had no reason to be taking an antidepressant so, that included a lot of nondepressed women to put it crudely. And so to get around that they said let's just use as the comparison group those who had a diagnosis of depression. I'm speaking somewhat colloquially you can be precise and that removed a few of the Zoloft users what dropped from 14,000 to about 11,000, but it of course reduced a huge number of the blue group and they had a -- but still had a 180,000 left and now you can do the same comparison, what's the risk of the cardiac birth defect in the red versus the blue, the Zoloft exposed to the unexposed and now it's dropped a small amount to 1.16.

Q. And what do you mean by its dropped a small amount?

A. Well, it went, I think it was 1.27 before and it dropped to 1.16 so that's important particularly since we close to one in this study for a number of reasons we could discuss. So, it's important from a statistician, you're not able if you look at the 1.27 could you flash back to that slide for a second?

So, the lower banned of that 1.07, 1.52, 1.16 is right in the mix there. It's hard to tell the difference between 1.16 and 1.27, but nevertheless it's moved it towards one

that's where you ought to focus on, because that's a more homogeneous, these groups are more homogeneous now, they all have a diagnosis of depression as compared to the one in the previous slide.

Q. So, now, in the Huybrechts study, Dr. Jewell there's an additional I think two levels of adjustment based upon something called propensity score analysis; is that right?

A. Correct. That's correct.

Q. Can you describe for us what that means a propensity score analysis?

A. Well, that's complicated.

Q. Okay.

A. It's basically a method for trying to balance so that you would want, suppose smoking was a little higher in the red group than the blue group and you were worried that that's distorting the rate comparison I just talked about, so you now adjust, you really re-weight the data to make the fraction of smokers in the both group the same by re-weighting basically counting down-weighting the numbers smokers in one group and up-weighting them in the other, basically.

And you do that for a whole bunch of [indiscernible] they actually did it for over 200 characteristics of the women and you need some massive software and some coding and you -- but the basic goal is to do the same, to balance the red and blue group on everything else that you know about. And so

they did that and then you compare the rate now you're re-weighted the data so the rates change a little bit when you've down-weighted some women in the red group or up-weighted in the blue you're down-weighting the weight you put on that outcome also and you get a new comparison and that's what led to their final reported result and it dropped I think to, well, I'm not sure it's here, but 1.09 as my memory. So, it dropped again, just a small fraction closer to one.

Q. Now, Dr. Jewell, you did something at the stage of depression restricted, well, first of all, does Huybrechts supply the various variables and data that they used to create their propensity score analysis?

A. No. They don't allow you to do that. They tell you about it, but you'd have to have the raw data to be able to change it yourself or re-weight things a little bit differently. All of these adjustment for confounding little differences that's why I'm not particularly caring about that odds are that they [indiscernible] odds ratio being 1.0 or .98 because you put in another variable it will make a little shimmy to it and change it slightly, but I can't do that even if I wanted to with the Huybrechts data because you need the original massive database and they aren't willing to release it for good reason.

Q. Now, so at the level where you did have information Dr.

Jewell, you did something with this group of women in this slide that would show appeared depression restricted, 11,059. You did an analysis of those women in particular; is that right?

A. Yes, because the Huybrechts team provided additional data on that red group that they themselves have deliberately collected to try and refine the measure of exposure by counting how many fills of prescriptions they had during the first trimester.

Q. So, if I'm understanding correctly Huybrechts original data in the paper that we did pull up and that's in the Court's binder actually provides data about these 11,000 women that include a particular whether or not these women filled a prescription during the first trimester or filled two or more prescriptions during the first trimester or filled none; is that correct?

A. So, that is correct. So we just talked about adjusting for confounding. We want to make the red women and the blue women as alike as possible on all these other factors and then they're only different, because one was exposed to Zoloft and the other was a depressed woman, but not exposed even to any SSRI.

But the blue women are -- the whole point of that re-weighting with having to measure 200 variables is because you don't believe the red women are really alike enough so

what --

Q. You don't believe the red and the blue in this analysis are alike enough?

A. Well, that's the point of doing that additional adjustment that took her from 1.16 to 1.09 that was that whole massive 200 variable stratification stuff, propensity score analysis. And okay, that's complicated you can't reproduce it's difficult and it's done because of this concern that the red and the blue women are still subtly different on other characteristics, I understand that.

But why Huybrechts themselves provided the data is you could take the red group and now compare red with red. In other words two different kinds of Zoloft women. So, you're taking the --

Q. [Indiscernible] that red?

A. Yes, it's all read in this case going back to the previous slide. These are the -- if you add these up these are the original 11,059 so if you flash forward again, take those 11,059, now they're all read and they're all homogeneous to the extent they were all suffering from a diagnosis of depression. They were all taking Zoloft to treat the depression. They all got pregnant, but then as we talked about Jimenez-Solem some of them behaved differently during the first trimester with regard to their Zoloft prescription than others and that's what this breakdown of this 11,000 is.

Some of them didn't refill the prescription.

Now, where he just talked about to be exposed in the Huybrechts data, you might have had a pill that took you one day into conception. That would count you as exposed, or two or three whatever.

If you didn't fill a prescription at that point you were really not very exposed if you only had wills for one day, in fact we don't even know if you took them, but, even if you did, according to your prescription you could only have taken them for one day, because we didn't refill the prescription.

So the red group here now is a group that is much less likely to be as heavily exposed to Zoloft to doing the critical first trimester than the blue group. We don't know for sure, no one does who took what on which day but as I just pointed out if you were exposed because you had pills that took you day beyond conception and then never fill the prescription you'll be in the red group.

If you were a woman who had pills on the prescription prior to conception that took you one or two days into pregnancy and then you refilled you'll be in the blue group. So this provides a very natural counterfactual experiment, meaning these women were all identical, we couldn't separate them up to the moment of conception. They were depressed, they got pregnant, they were on Zoloft, they were on Zoloft

enough most recently they had enough pills to take them up until conception and a little bit beyond and then some of the women stopped filling the prescription and some didn't and that's the red and the blue comparison here. As I said, shark of randomization which we can't do. This is actually the most natural homogeneous comparison possible. We're not bringing in women who never saw Zoloft. We're not bringing in women who aren't depressed. We're not bringing in women from different countries. These women were all indistinguishable as a group until the moment they conceived and then for whatever reason something happened to separate them into red and blue. Most of them kept taking Zoloft, 8,600, but 2,442 and Huybrechts had separated out the data by these two categories in her data supplement in her paper.

Q. So what did you find Dr. Jewell when you ran this analysis?

A. Right. So this is a natural experiment. I always was curious because of doing this is the big point of Jimenez-Solem to do this natural comparison. He is the only other study I have the data to do this and you see a 90 percent, 87 percent to be precise increase in risk of cardiac defects in the blue group compared to the red group.

Q. Now --

A. So, the ones who continued taking it to the paused group and that's statistically significant. So that's not

explainable by chance. Now, this isn't torching. I didn't try 15 of these kinds of analyses. I didn't try of these kinds of analyses. I went from Jimenez-Solem I read Huybrechts I said you can do exactly the same thing as Jimenez-Solem did in Huybrechts, she provided the data to do this crudely and I did it. In fact that's the whole point of trying to refine the exposure and get groups of women and restrict a depression diagnosed women and then now to all women, all these women taking Zoloft that's the whole point of this analysis in it's as powerful a comparison as you can get, but it doesn't present itself very often, only in Jimenez-Solem with its limitations and in Huybrechts with its limitation are you allowed to do this kind of comparison.

Q. Now, Dr. Jewell, because exposure in Huybrechts means any pill that could, that overlaps with the first trimester, doesn't that mean that the woman who fills a prescription for Zoloft in the second month of her pregnancy ends up being "exposed" even under your analysis?

A. Yes, Mr. Cheffo made a big point this morning that this is a ludicrous comparison, maybe I'm overstating how he didn't like it, but by [indiscernible] these are both exposed. How can you compare exposed and exposed?

But I just point out they're not the same, one was continually refilling. Some women in the blue group would fill their prescription twice in the first trimester, some

only once. The one in the red never refilled it so, they're not the same. Now, he's right, some of those red people may very well be exposed and they might have been exposed at the critical moment of organogenesis. I accept that, I can't separate them out because no one was watching the women, I don't know that. If they were hypothetically exposed at the critical moment, what would that do to the rate of defect if there is an association? It would increase the rate; right? Because they'd be exposed to Zoloft at the critical moment. I've treated them in this analysis as unexposed.

If in fact they weren't exposed the 1.87 would go up. So if Mr. Cheffo would provide the information from Dr. Huybrechts we know it's impossible, but if we hypothetically get this information and get rid of those exposed people from the red group the odds ratio would be even higher, that's obvious and demonstrable.

So, I don't have the same and in fact, I said this in the report and in my deposition, that's why that didn't bother me as much as you might think that the red group might have some people who in fact took the separate, you know, they took pills longer than their prescription or they took them through the critical period, because it would only -- if I could get rid of them, it would only make it worse for Zoloft, so I'm being conservative here in this comparison if that was really their concern.

Q. So, Dr. Givens filed a report in this case, Dr. Jewell where he took some umbrage of the fact that first he agreed that this information is available in the Huybrechts study and this is a quote on the slide called Givens Challenge his paragraph 23 from his report said, "Professor Jewell has also not conducted his new analysis separately for women who filled two or more prescriptions," and that data is available in the same table and he describes it this would guarantee a refilled prescription during the first trimester. Do you think that's a good idea?

A. Well, it's fine to do that. I did it by anyone filling at least one, so I included people who only had one filled prescription in my blue group, the exposed group. He wanted me to go even further by saying if you're really going to do this the way he would like, he would like it to be more definitive in the blue group now and say let's get rid of the ones who only filled one prescription, let's take those who fill two or more -- to use his words guaranteed that their exposed.

Q. What's that look like?

A. So, I did it of course after I saw his report and it's an identical result to within .01, as 1.88. This is now in the blue group only looking at women who had prescriptions leading up to pregnancy, filled a prescription at least twice during the first trimester as compared to the red women are

still the same. They had drugs taking them up until conception but never filled a prescription so it's a more clean if you want, I've got rid of the middle group here where they're, you know, what's with those women that filled it once? And it's exactly the same result and even though I've thrown away some data here it's still significant. So, there's clearly a difference between those women, blue women the continued users compared to the women who are not filling prescriptions during the first trimester.

Q. So if we say that this, the blue women have filled two prescriptions in roughly a three-month period for the first trimester of pregnancy that would exclude then women who might otherwise have been in that group who didn't start their prescription until third month of pregnancy; right?

A. Yes, I think his concern, I haven't talked with him and he didn't explain it, but I think the concern might be that you would have a woman who had enough pills to take into the first few days postconception stopped and then filled a prescription a day before the end of the first trimester that would count in my first analysis as one prescription fill and that particular woman wouldn't have been exposed in the blue group even though I called her as more exposed. So, you said let's get rid of that by doing the two. And of course it gives you the same results, by the way the bias in that if one could remove from the analysis from the blue group women who

weren't truly exposed if there are really some women in that blue group who are not exposed and therefore at no risk it would only make the odds ratio bigger again. So these two concerns, these two challenges haven't been thought through. They actually would make it worse for Pfizer if those women were to be removed if we had more definitive data. So, I'm actually here being conservative with regard to this comparison.

Q. Now, to be fair Dr. Jewell, one of the criticisms of this analysis that we've heard from Pfizer is well, look Dr. Jewell we know there was a stratification of the crude and then they did a depression restricted group which is what you used for your data, but in the Huybrechts study they went down to propensity scoring and even two levels of that if you look at the supplement. This is not adjusted for those propensity scoring Dr. Jewell.

A. That is correct and that was the content of -- I discussed that at deposition, I couldn't do that because that would require me having the individual level data to have those variables to adjust for confounding at this point and Dr. Huybrechts and I in that brief e-mail exchange discussed this point and as I said, that's exactly why I need the data or have you guys run it for me if you want, I don't have the data in my hands I've done that before where someone will do it and I'm blinding, so there's no confidentiality. That's to

take care of that loose end, there's no adjustment for that additional confounding.

Two reasons I'm not -- I'd love to do it and I asked if I could and I've been, I think the last message from Krista Huybrechts was we're talking to our team about doing it for me, basically. So, I think she -- maybe that will come through, but why I'm not that -- I'd live to see the results why I'm thought that concerned to be that I'm misrepresenting anything here is one, is already these groups of women are much more homogeneous than the group we -- that they applied their propensity score to which included all women who had a depression diagnosis but had never seen a -- never come within an arm's length of an SSRI. So that's a more heterogeneous group but this group as I said these women were indistinguishable until the moment conception happened so there's likely far less difference between them than there would be in more heterogeneous groups. So, confounding won't make as big a difference and two, when you look at where there was more heterogeneity in the original mean or comparison of Huybrechts and they did that adjustment for confounding it reduced as you saw from 1.16, I think the 1.09. So, it changed the odds ratio by .07 or a few percent.

Now, if you applied said well, we don't know how much difference adjusting for confounding of these results would be anyone like me, any statistician is going to say, well, what

plausible, what kind -- given that I can't do it today what's plausible about the impact of confounding? I saw 4 percent effective confounding on the more heterogeneous case even if I applied the 4 percent to the 1.88 that's going to reduce it to 1.7, it's not going to reduce it to one, so this is just not plausible that confounding can explain this result given what we know about confounding in the entire data set that Huybrechts is provided and when you see the size of the effect here.

Q. So, Dr. Jewell to summarize on this slide entitled Givens Huybrechts Givens Reanalysis you found it statistically significant one point I think you'd say that's 1.9 or 1.88 increased risk in a group of women who were all on Zoloft at or around conception and all were depressed?

A. And all took -- yes. That's the whole group. They're all exactly the same at that point still. And then the blue group is those women went on to take -- fill prescriptions at least twice during the first trimester and the red group never filled a prescription.

Q. Did you create any of this data or was this data available? Did you [indiscernible]?

A. This data for the crude comparison is all available in the Huybrechts paper.

Q. And Huybrechts supplied it because part of what she discussed was we looked at -- we were concerned about our pill

counting methodology so we also tested that to see if taking one or more prescriptions during the first trimester as compared to our pill count it changed the results or taking two or more it even changed the results. So, this was part of their methodology.

A. It was part of their methodology, part of what they called their sensitivity analysis. There they focused on the blue groups here and compared them back to the unmedicated depressed women and it never occurred to them to do this comparison and when I was interacting with Dr. Huybrechts since earlier this year when she asked me to explain it, she got it finally and said yes, I think this is an interesting possibility. Let me get back to my team with it I'll be able to give you what you want which is the confounded adjusted version of what I did.

Q. And that's what she wrote in this e-mail; is that right?

A. Right. Yes. This is where she says, it's an interesting possibility our agreement doesn't permit to share the data with me why don't I speak with others and assess what we can do from our end. So I'm very hopeful they'll do that so I don't have to have this discussion about that 1.9 how much it really changed, let's just do the analysis.

Q. And what does she mean when she says it will be surprising or what is your understanding? The scientist, what

she's trying to communicate when she says it'd be surprising and this an e-mail from Dr. Huybrechts to you dated July 12, 2015. What's your appreciation of what she means when she says, it would be surprising for post hoc subgroup analysis such as the one you described to supersede all other prespecified analysis?

A. I actually didn't quite know what she meant by that, because she's misinterpreting what I was trying to do. There's an e-mail I sent immediately back to this one saying I have no intention of the superseding, your main results they are very interesting in themselves, they're all in my report, but this is a particularly interesting comparison and now I'm using not just my words, I'm using her words too. And it's interesting in a way I tried to explain to her because we have the results from Jimenez-Solem that aren't identical, but are trying to achieve essentially the same thing and they're the only two papers that allow us to get this very homogeneous comparison of Zoloft users depending on and comparing whether those who kept using it versus those who didn't.

Q. And that's what led you to do a meta-analysis of your own in your report; is that right?

A. Right. So the Jimenez-Solem analysis that we discussed earlier comparing their non-paused with the paused group of users was not statistically significant. It's still a sizable risk there's some reason to believe that may be understated

for reasons we've discussed already, but now we have Huybrechts. They're not identical. The other hand they're giving very similar pictures and if you of course put the two together it becomes now highly significant. Your now comparing in both cases people who took Zoloft by some definition throughout the first trimester with women who used to take Zoloft, but didn't given that proviso that Jimenez-Solem didn't -- couldn't quite give us that, but -- and so that's -- these are very similar results there's no heterogeneity here as you can see at all, so.

Q. And Dr. Jewell, we're looking at a slide that's Meta-Analysis Huybrechts and Jimenez-Solem and this describes the results and of combining in a meta-analysis context Huybrechts and Jimenez-Solem and the output which shows an increased risk of 1.8.

A. Right and I just need to stress again, this is in some sense the most natural comparison you would want, because all of the other studies we struggle with this who do we compare the SSRI users or Zoloft user specifically with? What's the right comparison group of women and that has been a big discussion in the literature including this whole that spawned this idea of confounding by indication.

So, both Jimenez-Solem and Huybrechts went and some others have gone one step towards resolving that by looking only at women with a diagnosis depression. I just took it a

very natural next step. That's why I thought it was interesting Pfizer urged me to publish this at trial. I've now tried to, but I can't, I don't feel comfortable doing this without -- I don't want to use Huybrechts' data and take it and published on my own. I prefer to have something as a joint publication.

THE COURT: Did I just hear right? Did you say Pfizer urged you to publish?

THE WITNESS: Well, they asked me why, if I thought it was so important why had I not asked to, why had I not written a paper so, I took their advice and said, yes, better do that. And that's what led to the e-mail to Huybrechts.

BY MR. ZONIES:

Q. You mentioned heterogeneity when you did this meta-analysis, Dr. Jewell was there heterogeneity that was evident as there was in miles and the other meta-analysis?

A. No. You could see that with the naked eye there. That 1.9 and 1.6 are given those confidence intervals you can't statistically distinguish between them so there is no evident heterogeneity here. So it makes sense to average in this case, I mean, here, you're averaging two, so you can either, whatever your fancy you can look at the two results separately if you want to say well, the Huybrechts is just significant I forgot what the P value was flora bound is 1.1. It's a little

bit more significant when you combine the two studies which is the point of meta-analysis, of course.

Q. And how many studies were in wang; do you recall?

A. In wang?

Q. Yes.

A. Three.

Q. Throughout your report Dr. Jewell, do you also spend a significant amount of time looking at other confounding factors and potential biases that may have effected these results beyond just detection bias and confounded by indication?

A. Yes. As I've already alluded to several times there's this concern that women in the comparison group may not be identical to women in the exposed group on a bunch of behavioral characteristics, like smoking, obesity and so on and so -- and this is an observational study so that it's not resolved by randomization of exposure as it would be in a drug trial. So, here you've to work hard at this and most of the authors tried depending on the quality of their data to do this at various stages so we do see how much the results change as these authors were wrestling with this adjustment for confounding.

Q. And the last one is other unknown factors residual confounding, can you describe what residual or unmeasured confounding means?

A. Well, that essentially means when an author or team, research team has tried everything they have measurements on to get rid of these possible differences between the exposed and unexposed women. They might still worry because there's no randomization that the difference in the risks let's say for cardiac defects that still exists might be due to something we didn't measure and that's what people refer to as unknown or residual confounding.

And of course any observational study from the beginning of time always has the possibility that there's some unknown factor. But we're not completely powerless as statisticians, because as I indicated earlier, we know for example to get rid of that 1.9, you need a fairly sizable effect on the outcome cardiac defects that you don't know about to be able to explain it away. So, there's concern about it, but there's also -- we're not completely hamstrung by we're able to think about what the potential impact might be of an unknown confounder.

Q. And we'll just touch on Bradford Hill factors, Dr. Jewell you mentioned that strength of association was the first factor. What did your review of the data show with regard to strength of the association?

A. Well, strength of the association is the idea it's related to confounding that the stronger the effect you see the less chance that something unknown could be -- dealing

with. So if you have an odds ratio of 1.1 or a [indiscernible] risk of 1.1 then it's not so hard for an unknown confounder to wipe that out, because it only needs to have relative risk of 1.1 and there might be things around that raise the risk of a cardiac defect by 10 percent that we don't yet know about. A lot of cardiac defects we don't the cause so that's possible, but when you get up to two or three, it becomes much less likely that something you don't know about is completely going to explain away that result. So, that's why in Hill's criteria strength of association is important. It is particularly if you're dealing with making a policy change. You want to be reasonably sure. So the stronger the association the less chance it's spurious.

Q. And again is this just oh, you look at the numbers and you go oh they are this? Or is it well, you know, this I feel is stronger the Huybrechts analysis outcome for a strength of association is stronger because I've done a very deep dive on trying to remove all of those confounding factors?

A. Well, as I said, the advantage of that analysis is that you've really done amongst a very homogeneous group of women. The others don't have that advantage but their very strong effects. And so there are some very strong effects here that would concern any of us.

Q. Consistency of effects is the second Bradford Hill factor and on the screen is the bond paper discussing at least

those investigators just conclusions that there were consistent effects for an increased risk of cardiac birth defects with the use of Zoloft. Is that consistent with your conclusion as well?

A. Yes, the idea here is if you've done an observational study and you see an effect and you get alarmed and I'll tell you it happens, it crosses my desk quite a lot, the next -- but then there are other studies and there's no effect if you've just seen it once that makes you suspicious.

So the idea of consistency here is not that you expect every single result to just come in on the nail, everything is the same, I wish that it were, I would be -- have a much simpler job. They come -- they bounce around a bit and we've seen that in the visual representation.

The idea though is, is it just one stud y? That's obviously no. Is it just two? No. Is it just three in this case? No. And so, we've seen enough here to say this is not just one spurious thing that happened by chance in one population. That's just -- I'm not just relying on that myself, other authors constantly state in their paper is this result consistent with what we know so far. And you've picked out Bamm [ph] which is in the United Kingdom, I think. Here's -- well, Pfizer's own epidemiologist characterizes this consistently positive. They don't mean every study gave the same result is you can see in what they were report in this

paragraph, that's not true but they do see enough to say there's something consistent going on, it's a tricky English word, but consistent is is it's not something that just happened once by chance coincidentally.

Q. And we're looking at slide, Dr. Jewell, entitled consistency Pfizer epidemiologists agree and it's a Pfizer document in 2014 where Pfizer's scientists conclude that there's a consistency insistently positive association for searching the exposure and cardio vascular defects especially septal defects. And that's something you agree with; correct?

A. That's consistent with my opinions, yes.

Q. Did you reach your conclusion about this even before you saw this?

A. No. Actually I didn't see this document until the Frye [ph] hearing.

Q. So, it wasn't like you relied upon this to reach your conclusion, you just saw it's in agreement with your conclusions?

A. That is correct. And to me it's not so much important what their opinions are it's the methodology that's consistent that they were doing the same kind of thing, albeit in a smaller set of studies as was pointed out this morning by Mr. Cheffo, but the methodology was still the same. You take the studies, you look at them and you see if there's a consistent positive association.

Q. And the next slide is again the flores plot that we will have enlarged by tomorrow showing all of their results and as you've said before part of what you do is look at this to see the symmetry I think or the asymmetry, I'm sorry --

A. Well, first of all it's -- I was going to say it's effective visually to see all the results on one page. We've just determined that it's not that effective. So, I understand that and I can't read it myself and I've got contacts in. But it does in some sense the blurring may actually be valuable here, because it's saying what you really want to get out of this is not the numbers, you know, the fine detail on the left-hand side, the thick solid line is one. The number, the relative risks are odds ratio run on the bottom axis from .1 to 10 equally, because there's an asymmetry on the axis, so you may -- .1 is the same distance from 1 is 10. One protects you tenfold, one is tenfold bad, so that it's balanced that way, so we're not distorting it by not using a log scale.

And then just look at them and say, well, what's your -- is this completely down the middle? Is it asymmetric to the right or to the left? Now, we can't do this in the courtroom, what I always do is if you're struggling, I don't think there's a single person in this room that really is struggling with the IP address symmetry, it's pretty obvious.

But if you're really are still struggling just turn it

upside-down and see if you still feel the same. If you can't tell the difference whether there's more stuff pushed to the right, meaning an increased risk than there is to the left meaning a decreased risk, if the relative risk is one everything should be sort of balanced on the each side; right by chance? If you can't tell the difference between this when you look at it this way and turn it upside-down then I can sort of say, well, it didn't work for you, you see no difference. When I do that, when I turn it upside-down I see that the asymmetry has shifted to the left by my eyes and so that's what this really represents in a visual way there is this push to the right rather than the left. Is it absolutely the same everything single study, I never have seen that and this is tough but it's not what you expect.

Q. And to be fair and we'll see it when it's clearer. This includes repetition of findings within studies.

A. It does.

Q. In other words certain studies like Jimenez or Coleman [ph] I think Coleman shows up the most on here; right?

A. Well, some of the papers reported more subgroups than others, but you can do this if you want by taking just the all cardiac outcomes.

Q. So, we talked about this book by Rothman and Greenland called Modern Epidemiology, are you familiar with this book?

A. Yes, and I know two of the authors quite well.

Q. And in this book there's this quote about consistency, because I mean, I suppose I'm a big concern that Pfizer is going to get up and say, well, all of these aren't significant you shouldn't look at those which frankly if we were looking at consistency and only looked at the significant findings well, heck, we could go home. But you actually look at all of them including the nonsignificant protective effects; correct?

A. Yes, they're important and if there's enough of them they're going to influence your opinion.

Q. And is it consistent your methodology with this discussed in this book in this [indiscernible] treatise that it sometimes claims that a literature or set of results is inconsistent simply because some results are statistically significant and some are not. This sort of evaluation is completely fallacious. Even if one accepts the use of significant testing methods. The results from a set of studies could all be identical even if many were significant and many were not et cetera. Is that consistent with your methodology and the methodology that you recognize as that in your science?

A. Yes, that's generally true. And one of the reasons is that significance is driven by more than what the truth is. It's driven by the size of the study and many other things, so you can have two studies that have really studying the same population that have the same results one is just not

significant because it's not powered enough, it's not big enough, but does that it doesn't provide any information? No. You should use all the information and significant results and nonsignificant results should both be assessed and you've got to weigh those in your coming to any reasonable scientific opinion.

Q. And if we were assessing consistency and relying solely upon significant results would there be an asymmetry in this?

A. Well, that -- yes, because there are no negative asymmetry. I would never believe, I would've been surprised if there were as Mr. Cheffo would also have been shocked if there had been a significant result on the left, because no one really believes that the drug is going to protect you from cardiac defects and if you just studied it enough it would show up. So, I understand why there isn't any. No one is surprised at that. But that's one of the reasons why it would've been unfair to me to Pfizer to ignore the ones that are somewhat negative. They're negative and they show no increased risk all albeit they have imprecision attached to them.

Another main reason why you shouldn't just focus on significance in balancing though studies, those studies are not all equally good. We haven't talked about that that much. We've mentioned it, oh, somebody said, Huybrechts is a great study or Jimenez-Solem is a great study, because it's so big

and has so many women. Big studies don't have any guarantee to get rid of bias. They get rid of -- they help you with precision; right? But they don't do anything about bias.

The largest presidential poll in the history of the United States protected Dewey would beat Truman; right? And it was a shock and it was the birth of modern polling, because people say we -- everyone said Dewey was ahead. It was a huge poll, but it was biased, because it was subscribers to the literary digest and they were -- had a particular political bend.

So, big polls if their biased are useless because even though they get very precise results they don't necessarily help you. So, we've talking about these studies and the visualization fails in some regard, because it treats each line as somehow equally valid other than the sort of confidence interval. But that's not true, there are some very good studies in here in my view when you analyze them with a statistical assessment of bias and there are some rather poor studies and we -- I could now if you want at some point or if Pfizer wants I can tell you which ones I think are poor and why I believe that, what's the objective criteria. But that's the danger of just putting them all on an equal footing either in a plot like this it's dangerous to do -- to read something quantitative in this and it's also dangerous if you just throw them all in a meta-analysis, because you're not dealing with

that at all.
Q. Well, and that's a point and that I want to make it perhaps is my mistake that I want to ensure that everyone understands, Dr. Jewell, which is you didn't assess consistency with a flores plot; did you?
A. No.
Q. When you were doing your analysis?
A. No. The flores plot is merely just a way to get all of the studies on the page so we can talk about them and find them quickly.
Q. And in fact in your paper you go into great detail about which studies you find strong and why and which studies you have concerns about and how you would individually assess those and how you did in fact individually assess those?
A. Yes, and that's a very important topic we really haven't addressed so far today.
Q. Specificity or the short-term shorthand term for this that I don't mind saying is lumping of endpoints and one of the primary concerns in the original doubt we're hearing in this case was looking at an endpoint called all congenital malformations, that's not something you were comfort doing was it Dr. Jewell? Looking at that endpoint.
A. No. I would've, I always prefer to get as specific a charge as possible, specific in endpoint and one can do it looking at all malformations but there's a lot more

heterogeneity in that clumping than there would be in a subcategory.

Q. And as you described very carefully in your report, you didn't as we said a number of times -- this quote from your report at page 17, you didn't independently modify or poll the endpoints; correct?

A. No.

Q. And you go into detail in your report again, Dr. Jewell about how you relied upon the experts in the field including as we see up on the screen here the EUROCAT European Surveillance of Congenital Anomalies Association in part on their choice of classification and grouping and even they say lumping together heterogeneous sets of anomalies; is that right?

A. Yes. This is -- EUROCAT and their summarizing I think I actually probably used this quote, because they summarize while the conundrum that I mentioned earlier when you're trying to group outcomes that you know are not identical as I said about lung cancer. Pathologists know they are not identical but if you get down to two fine a grid, that's their B there, if you split so finely there are so few cases of this very fine subgroup that you can't distinguish, there's not enough information, in fact there would be a tactic to use if you wanted to get to no statistical significance just get down to a very fine subgroup. I'll guarantee you no statistical

significance. Any study you want, give me now in 30 minutes I will make a finer subcategory and the statistical significance will go away. Would you believe this? No. Right, that's just Machiavellian.

The other side is, oh, if you put too many together, let's go the opposite direction. If you put too many together and you said let's do all congenital abnormalities now you really are talking about different organs of the body that are effected at different times during pregnancy and you may be diluting any specific association so every expert has to -- what is it they say strike the balance here. And I would've been faced with that where I left on my own, the balance I struck was to use the balance struck by the original investigators. So I didn't try and add something into cardiac or I didn't do anything they didn't do.

Q. And the last criteria or last viewpoint of Bradford Hill biological plausibility. And Dr. Jewell do you agree with this quote, I'm showing a quote from the reference manual on scientific evidence 3rd at 604 that says when biological plausibility consists it lends credence to an inference of causality, because as it says, just above that biological plausibility is not an easy criterion to use and depends upon existing knowledge about the mechanism. So, when it does exist, it actually means something good.

A. It's important, in fact, this is of course the one area

I'm not an expert as had been pointed out. I'm not a cardiologist. I'm not an expert on perinatal birth defects. I'm a statistician. So, I'm an expert at looking data that other people have defined and categorized and put together.

They're not so good at what I do and I'm not so good at what they do. So, when it comes to biological plausibility, I personally cannot, I do not study the biological mechanism during organogenesis. I have to rely on other biologist in this case. I have consistency, that's a statistical issue. Strength of an association statistical issue. Hill was a statistician. Biological plausibility does it make sense if the flowers in your garden are red that you suffer from a birth defect? Probably not. There's no biological plausibility unless somebody comes forward, we've said oh there's an allergen in it, and then there's a biological mechanism, but you have -- if there's no biological mechanism known to the scientist who are experts that weakens the case I think of causality substantially.

Q. And here you've found, you're relying upon the experts in this case that indeed there was a biologically plausible mechanism?

A. Yes, I think there's a bunch of evidence that's not evidence that I have to take other expert's opinions.

Q. Well, and one of the issues Dr. Jewell is -- and is this about the difference between SSRI drugs and your

biological mechanisms in fact. So when you did the Huybrechts analysis, subanalysis for Zoloft you came up with a statistically significant [indiscernible] almost double [indiscernible]. In your Prozac work did you also do a similar analysis with the Prozac data from Huybrechts?

A. Yes, I did. I did exactly the same analysis for Prozac, I actually did it for Paxil too just to get a sense of what the data was saying. I testified to that when I was deposed on Prozac, I presume Pfizer has a copy of that. I testified it when Pfizer deposed me on a second deposition, so I have said many times that I have done it, it's not true that I have never done it for Prozac.

It's not that I haven't done it, Prozac is a different drug as is Paxil, so it's just different. But I have done the calculations and they've seen -- everyone in this room whose interested has seen my analysis of those results.

Q. But Dr. Jewell, you didn't include those results in the Prozac report and Pfizer in their opening said oh, well, that's because he didn't like those results. Is that why you didn't include those in your Prozac report?

A. No. And those results by the way were in the e-mails to Krista Huybrechts, so I know everyone has them now where I was trying to point out that let's do the -- excuse me -- let's do this analysis of like with like that we talked about, the pause and the non-pause. I didn't want to write

necessarily only Zoloft and then be criticized why Zoloft. Let's do it for all of the ones that we have data which does include Prozac and Paxil. And those results are in those e-mails and then of course you have to address in the case is the half-life of drug itself and Zoloft has a shorter half-life than Paxil or Prozac, so when you stop taking pills in a sense you don't fill the prescription, we don't know what the woman is doing, but if there's any left it's not going to be as much in their system a week later or two week later as it would be for Paxil and Prozac. That limits the value, you know, we remember just struggling, I've been criticized already today about, well, they're all exposed, you can't really tell the difference well, the best chance of telling a difference with Zoloft and a much more difficult task with Prozac and Paxil and I was aware of that because of the half-life properties of the drugs.

Q. How do you feel about being called a situational scientist?

A. Well, people call me lots of things. You know in some senses all science is situational. You don't treat dogs and cats the same way. If they have a thrombosis you treat the dog differently from a cat. It's situational, it makes sense to me. Well, I'm a statistician. When you see the same question let's say in this case Prozac or Zoloft do they cause cardiac birth defects, it's the same question, but their

different drugs that -- what's important to me as a statistician is what does the data say, so I have to look at the data separately. I'm not going to say oh, Prozac causes birth defects as I think it does in Prozac or vice versa. Or even if I believed one of them didn't cause it, I have to -- I'm trained to tell you what the data says and not play games and not let it be influenced by my opinions or not, I have no interest personally on taking one side of this, partly because I'm involved in many -- one of the statistician's job is to have to be involved in many of these questions simultaneously. I've got about four in my head right now where I'm thinking does this drug cause this problem or not. And in one of them I think it doesn't I'm not going to tell you which one. And in three, I think --

Q. Well don't tell them.

A. Well, they don't care. It's not their drug. And three, I think it does, but I don't, I can't afford to have a stake in each of these four decisions, I just take the data, what I enjoy and what I find satisfying as being trained to do well is to look at the data and tell you on either side this is what the data is saying. That's situational in the sense that the data is different for different situations. But I certainly don't appreciate the idea that I would just change my opinion and the data says the same just because there was some vested interest. I have no entrust in the outcome

whatsoever from that perspective.

Q. Were you ever hiding from the difference between those two?

A. No. And in fact in my deposition on Zoloft I begged the deposing attorney to be able to discuss with him Prozac but he would not discuss Prozac with me during my deposition.

Q. Thank you for your time today Dr. Jewell.

MR. ZONIES: Your Honor.

THE COURT: That concludes your examination and we'll start with cross-examination in the morning. We'll start at 10:00 tomorrow and I think I don't need a board I'm just asking for a 11 by 14 size.

MR. ZONIES: Understood.

THE COURT: Okay? So it should be something easy to do. All right?

MR. ZONIES: Appreciate that. Thank you.

THE COURT: Thank you. Have a good evening.

- - -

(Whereupon, the proceeding was concluded at 5:21 p.m.)

- - -

C E R T I F I C A T E

I do hereby certify that the aforesaid hearing was transcribed by me from an audio recording to the best of my ability; and that I am neither of counsel nor kin to any party in said action, nor interested in the outcome thereof.

WITNESS my hand and official seal this _____ day of _____, 2015.

Janine Thomas
Notary Public